SYNTHESIS AND STRUCTURAL STUDIES OF CYCLIC Py-Im POLYAMIDES

Thesis by

David Michael Chenoweth

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended January 15, 2009)

...for Kimberly, my love...

Acknowledgments

This thesis is dedicated to my wife Kim for her constant love, support, and friendship through the years. I can't thank her enough for her love and support, which has guided me through our time at Caltech. Our scientific collaborations and discussions have been another source of constant enjoyment for me. This thesis would not have been possible without her help and support and I am filled with excitement as I think about the arrival of our first child and our future together.

I would like to thank my Ph.D. committee Peter Dervan, Linda Hsieh-Wilson, Doug Rees, and Brian Stoltz for all the time and effort they have invested in me while at Caltech. They have all been fantastic role models and their dedication and passion for science has been truly inspiring. I would like to express my sincere gratitude to my advisor Peter B. Dervan for his guidance, support, and constant willingness to let me pursue all my crazy ideas and projects while in his group. I have truly valued his guidance and mentoring as I navigated my way through research and I thank him for the privilege to work in his research group. I sincerely thank the chair of my committee Linda Hsieh-Wilson for all of her guidance and support during graduate school. Teaching Chem 145 with her has been one of the highlights of my time at Caltech. I will miss the enjoyment of trying to make the best possible problem sets and test questions together. I think often times I learned more than the students during our endeavors. Doug Rees has been a fantastic mentor and I am grateful for his support throughout my Ph.D, his excitement about science, and his mentorship has profoundly influenced me and will continue to be an inspiration. My interactions with Doug and his research group have been one of the highlights of my graduate school experience at Caltech. I thank Brian Stoltz for serving on my committee. His passion for science, intellectual curiosity, and hard work are truly inspiring. The time he has devoted and his comments, honest criticisms, and mentoring during my proposal exams and thesis defense are greatly appreciated.

Many other professors have made an incredible impression on me while at Caltech and I thank them all. Bill Goddard has been a fantastic collaborator and mentor and his excitement about science is truly infectious. I will miss our discussions about science and collaborations on projects orthogonal to my Ph.D research. His generosity and willingness to give me access to his computers and group resources has been very valuable and is greatly appreciated. I thank Harry Gray for his mentorship and advice and for being another inspirational professor to me. I especially valued our discussions about research (fluorescent polyamides) and his willingness to share his group resources. Dennis Dougherty is thanked for his advice and research discussions (lone pair-π

interactions and electrostatic potential map analysis). Dennis has also been an inspirational professor and his Physical Organic Chemistry textbook has served as a constant source of enjoyment to me, often occupying my night stand and keeping me reading late into the night. Jack Roberts is thanked for our many discussions about NMR and his constant willingness to talk to students and help them with problems. Just walking down the hall between Church and Crellin every day and seeing Jack intensely pounding away on his computer has been very motivating. I always knew he was thinking deeply about science and this was very comforting for me.

I have had the privilege to work with many collaborators while at Caltech and I thank them all. Dan Harki has been a great friend and collaborator and I greatly enjoyed our time at Caltech and look forward to future collaborations. I would also like to thank Christian Dose for many great times in the lab and the rest of my collaborators that I have had the privilege of working with (Justin Cohen, John Phillips, Anne Viger, Mike Marques, Mike Brochu, and Julie P.). I also thank the rest of the Dervan group for making the lab a wonderful place.

I sincerely thank Mike Day and Jens Kaiser for putting up with my constant questions and nagging day in and day out during my adventures in crystallography. I thank both of them for their guidance and friendship during my time at Caltech. I will definitely remember the Friday night plasma party with Jens in the Rees group microwave and the many fun synchrotron trips. I thank Bill Scott at UCSB for help with crystallography software setup on my mac and the Chimera team at UCSF for their help with new software features and script writing.

I sincerely thank all of my friends and family outside Caltech for their constant love and support through the years: Mom, Dad, Dannielle, Ed and Sandy Shuler, and Kieth Shuler, Angela Harki, Sam Harki, and Tom Britton. We really miss all of you! Angela and Sam Harki are thanked for being great friends and listening to all the complaints during Grad school. We will definitely miss the two of you (and Dan too)! We will have to find some way to visit each year.

I would especially like to thank Tom Britton for his friendship and mentoring during my time at Eli Lilly. He has made a lasting impression on me and taught me more than I could imagine.

Abstract

The work presented in this thesis is focused on the molecular recognition of DNA by minor groove binding polyamides. Methods and strategies for the solution-phase synthesis of hairpin and cyclic pyrrole-imidazole polyamides are presented with optimized protocols requiring little to no chromatography. These synthetic strategies have led to the design of cyclic polyamides targeted to the androgen response element and are shown to be biologically active and cell permeable in cell culture experiments in addition their binding affinities rival that of most polyamide architectures. The structural elucidation of an α-amino-turn-linked cyclic polyamide is presented at 1.17 Å resolution providing insight into the detailed molecular recognition process and allosteric modulation responsible for the inhibition of transcription factor-DNA binding. Additionally, structural elucidation of a β-amino-turn-linked cyclic polyamide, highlighting the conformational differences compared to the α-amino-turn linked structure is presented. A structural basis for the inability of polyamides to bind dsRNA is also proposed based on biophysical, structural, and modeling data. In addition to these studies a new class of programmable oligomers targeting the DNA sequence 5'-WGGGGW-3' were shown to inhibit DNA binding of the Nf-kB transcription factor by EMSA gel shift. Compounds synthesized in this study were found to possess unique fluorescent properties with the ability to modulate their fluorescence by binding their targeted dsDNA, leading to sequence specific fluorescent detection reagents. Efforts toward the templated-assembly of polyamides using higher-order DNA structure (NCP) are also reported and the development of a new profluorescent class of heterocycle, which has the potential to be used as a chemical reporter of ligation events is described.

Table of Contents

List of Figures

CHAPTER 1: Introduction to Molecular Recognition of DNA

CHAPTER 2: Solution-Phase Synthesis of Pyrrole–Imidazole Polyamides

APPENDIX D: Peptoid Cell Uptake Studies

APPENDIX E: Cell Uptake Studies of 4G Targeting Polyamides

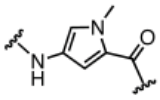
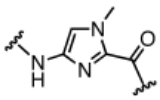
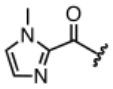
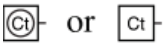
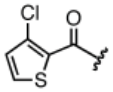
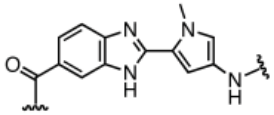
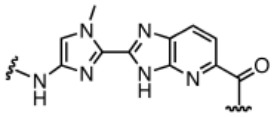
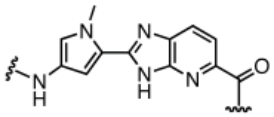
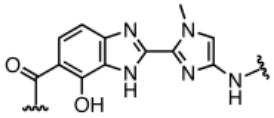
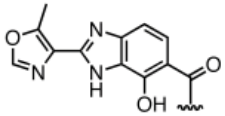
List of Schemes

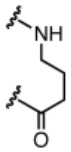
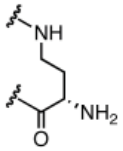
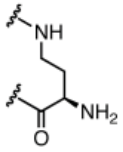
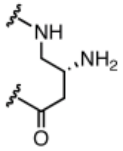
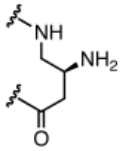
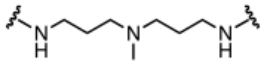
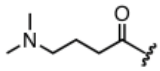
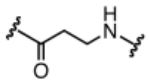
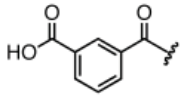
List of Tables

Nomenclature and Symbology

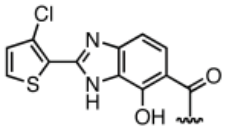
A	adenine
Å	angstrom
A•T	adenine Watson-Crick hydrogen bonded to thymine
Ac_2O	acetic anhydride
ADMET	absorption, distribution, metabolism, excretion, and toxicity
AP	activating protein
AR	androgen receptor
ARE	androgen response element
atm	atmosphere
Bi	benzimidazole
Boc	*tert*-butyloxycarbonyl
bp	base pair
ºC	degrees Celsius
C	cytosine
calc'd	calculated
Cbz	carbobenzyloxy
CCDC	Cambridge Crystallographic Data Centre
Ct	2-carboxy-3-chlorothiophene
Dbu	diazabicycloundecane
DCM	dichloromethane
DHT	dihydrotestosterone
DIEA	N,N-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
Dp	*N*,*N*-dimethylaminopropylamine
ds	double strand
Em	emission
ESI	electrospray ionization
Et	ethyl
Ex	excitation
FAB	fast-atom bombardment
Fmoc	fluorenylmethyloxycarbonyl
G	guanine
g	grams
G•C	guanine Watson-Crick hydrogen bonded to cytosine
GABA	gamma-aminobutyric acid
h	hour(s)
HBTU	2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HF	hartree fock
Hp	3-hydroxypyrrole
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry

hν	light
Hz	hydroxybenzimidazole
IC_{50}	median inhibition concentration (50%)
Im	*N*-methylimidazole
Ip	imidazopyridine
K_a	association constant
K_d	dissociation constant
λ	wavelength
m/*z*	mass to charge ratio
μ	micro
M	molar
max	maximum
MALDI	Matrix-assisted LASER desorption/ionization
min	minute(s)
mol	mole(s)
mmol	millimole(s)
MS	mass spectrometry
N	normal
N	A, T, G, or C
NCP	nucleosome core particle
No	oxazole
NOESY	nuclear Overhauser enhancement spectroscopy
PCR	polymerase chain reaction
Py-Im	pyrrole-imidazole
PNA	peptide nucleic acid
PSA	prostate specific antigen
Py-Im	pyrrole-imidazole
RT-PCR	reverse transcriptase PCR
Py	*N*-methylpyrrole
PyBOP	(benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
OBt	hydroxytriazole ester
R_f	retention factor
RNA	ribonucleic acid
RP-HPLC	reverse-phase high performance liquid chromatography
sat.	saturated
satd.	saturated
ss	single strand
T	thymine
TFA	trifluoroacetic acid
TMR	tetramethyl rhodamine
TO	thiazole orange
TOF	time-of-flight
U	uracil
UV	ultraviolet
Vis	visible
W	adenine or thymine

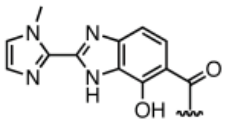
	-Py-	
	-Im-	
	Im-	
	Ct-	
	-BiPy-	
	-ImIp-	
	-PyIp-	
	-HzIm-	
	NoHz-	

	-γ-	
	-*(S)*$^{\alpha\text{-}H2N}$γ-	
	-*(R)*$^{\alpha\text{-}H2N}$γ-	
	-*(R)*$^{\beta\text{-}H2N}$γ-	
	-*(S)*$^{\beta\text{-}H2N}$γ-	
	-(+)-	
	+)- or Dp-	
 -β-	-β-	
 or IPA–	-)- or IPA-	

CtHz-



ImHz-



CtBi-